

基于改进灰色GM(1,1)模型的广西第一产业经济预测

石欣烨, 黄靖贵, 李炳炎, 隆璿瑶, 罗琳媚

南宁师范大学数学与统计学院, 广西 南宁

收稿日期: 2023年5月21日; 录用日期: 2023年6月22日; 发布日期: 2023年6月30日

摘要

建设农业强国, 是党中央着眼全面建成社会主义现代化强国做出的战略部署, 第一产业经济发展自然能为中国农业未来的发展趋势起到了指引作用。广西作为农业大区, 分析和预测第一产业生产总值(GDP)发展数据对预测地区经济发展有很大的参考作用。为此提出了一种基于改进灰色GM(1,1)模型, 对广西第一产业经济进行预测。为提高预测模型精度, 对2022年广西第一产业GDP季度原始数据进行数据频率的转换, 从低频数据转为高频数据, 提出了一种新的改进GM(1,1)预测模型, 描述其预测精度等级评价, 并利用2022年广西第一产业第四季度的月度数据对改进后的模型进行训练。

关键词

GM(1,1)模型, 灰色改进模型, 广西第一产业, 经济预测

Economic Forecasting of Guangxi Primary Industry Based on Improved Grey GM(1,1) Model

Xinye Shi, Jinggui Huang, Bingyan Li, Liyao Long, Linmei Luo

School of Mathematics and Statistics, Nanning Normal University, Nanning Guangxi

Received: May 21st, 2023; accepted: Jun. 22nd, 2023; published: Jun. 30th, 2023

Abstract

Building a strong agricultural country is a strategic deployment made by the Central Committee of

文章引用: 石欣烨, 黄靖贵, 李炳炎, 隆璿瑶, 罗琳媚. 基于改进灰色 GM(1,1)模型的广西第一产业经济预测[J]. 理论数学, 2023, 13(6): 1829-1840. DOI: 10.12677/pm.2023.136186

the Communist Party of China with the aim of comprehensively building a socialist modernized country. The development of the primary industry economy naturally plays a guiding role in the future development of Chinese agriculture. As Guangxi is a major agricultural region, analyzing and predicting the development data of the Gross Domestic Product (GDP) of the primary industry is of great reference value for predicting the economic development of the region. Therefore, a modified Grey GM(1,1) model is proposed to predict the primary industry economy in Guangxi. In order to improve the accuracy of the predictive model, the data frequency of the original quarterly GDP data for the primary industry in Guangxi in 2022 is transformed from low frequency to high frequency. A new modified GM(1,1) predictive model is proposed, and its prediction accuracy level is described and evaluated. The model is then trained using the monthly data for the fourth quarter of 2022 for the primary industry in Guangxi.

Keywords

GM(1,1) Model, Improved Grey Model, Guangxi Primary Industry, Economic Forecast

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 研究背景及意义

习近平总书记在党的二十大报告中明确提出，全面推进乡村振兴，加快建设农业强国，是党中央着眼全面建成社会主义现代化强国作出的战略部署。第一产业指农民和农业、林业、牧业、渔业等。农业是立国之本，强国之基，强国必先强农，农强方能国强。第一产业经济发展自然能为中国农业未来的发展趋势起到了指引作用，大力推进第一产业经济发展也是实现乡村振兴战略的必经之路。广西是农业大省(区)，以糖料蔗、蚕茧、木材、水果、秋冬菜、茉莉花等产量在全国居于领先地位。2022年，为全力打好农业稳产增收硬仗，使农业经济运行保持向好态势，广西大力实施乡村振兴战略，推动全区强一产促增长，壮乡乡村振兴事业实现高点开局。2022年，广西壮族自治区政府印发《广西科技强农八大工程实施方案》，旨在促进农业高质量发展。广西从农业大省转变为农业强省势在必行，在农业增产处持续发力，则需要科学预测分析第一产业经济值及未来发展趋势。

灰色预测是一种在样本量较小的情况下，可以确保精度较高的预测方法，它可以利用已知信息和未知信息来对系统进行预测。中国学者邓聚龙教授[1]在1982年创立灰色系统理论，该系统理论的基本思想上是对样本量较少、信息量不丰富的不确定系统进行研究；1992年李云贵等[2]对GM(1,1)预测模型做出了一定的改进，目前灰色系统理论的应用受到越来越多的关注，甚至在农业和灾害预测方面已展现出比传统预测方法的更大优越性。目前，许多国内外学者通过时间序列分析的方法来预测第一产业的总产值，例如吴晓明等[3]建立ARMAV模型对国内第一产业总产值进行预测，预测误差达3.05%；王妍[4]等运用GM(1,1)模型对山东省农业，总产值未来5年的发展情况进行预测，李勇等[5]使用灰色关联度分析来研究农业、林业、牧业和渔业对河南省第一产业增加值的影响程度。凌立文等[6]利用层级预测及其修正方法呈现了我国个别省份第一产业总产值未来两年的变动趋势；张仕清等[7]针对第一产业总产值及各部门子产值进行独立预测，提出模型结构的优化方案。然而，大部分学者研究中仅利用年度数据进行建模，也没有对GM(1,1)模型做出改进[8][9][10]，可能导致数据跨度过大导致误差率较高的问题；而且没有针对广西区的第一产业经济展开独立预测，对于广西区来说预测结果的可参考性及现实指导意义不是很强。因此，此研究方法与

模型适用于产业产值分析。科学预测第一产业经济的发展态势，对于进一步推进“三农”工作稳步争先，推动农业农村高质量发展，优化农业经济活力有着重要意义。本文选取 2022 年广西第一产业经济前三个季度数据，用二次插值法扩充广西第一产业经济月度数据，再利用扩充的数据预测未来 3 个月的经济值，为广西省制定未来第一产业发展策略提供一定的理论依据，推动全区强一产促增长。

2. 二次插值法原理

2.1. 二次插值法介绍

拉格朗日二次插值法的原理为在被插值函数曲线上任取相邻三个点作为插值节点，以这三个点作一条二次曲线来近似地拟合原曲线，此时该曲线又称为插值多项式。假设存在 n 个点 $(x_n, f(x_n))$ ，给定 $n+1$ 个插值节点 x_0, x_1, \dots, x_n 且 $x_0 < x_1 < \dots < x_n$ ，以及每个节点所对应的函数值 $f(x_0), f(x_1), \dots, f(x_n)$ 且 $f(x_0) < f(x_1) < \dots < f(x_n)$ 。构造插值多项式 $L_n(x)$ ，设计基函数 $l_n(x)$ ，使得 $L_n(x)$ 能够拟合 $f(x_n)$ 。即对 $k = 0, 1, 2, \dots, n$ 时都满足

$$L_n(x_k) = f(x_k)$$

则 n 次插值多项式可以表示为

$$L_n(x) = \sum_{k=0}^n f(x_k) l_k(x) \quad (1)$$

利用待定系数法求得基函数

$$l_n(x) = \prod_{\substack{i \leq j \leq n \\ i \neq j}} \frac{x - x_i}{x_i - x_j}, i = 0, 1, 2, \dots, n \quad (2)$$

$l_n(x)$ 的特点是，当 $x = x_j$ 时， $l_k(x_j) = 1$ ，当 $x = x_i \neq x_j$ 时， $l_k(x_i) = 0$ 。这样一来， $L_n(x)$ 就能穿过 $n+1$ 个 $(x_k, f(x_k))$ ，当 $n=2$ 时，即为二次插值公式。

2.2. 基于二次插值法的 2022 年广西第一产业经济计算

本文思路是，已知 2022 年广西第一产业经济前三季度值，记为 (x_0, y_0) ， (x_1, y_1) ， (x_2, y_2) ，且 $y_0 < y_1 < y_2$ ，过此三点进行二次插值。2022 年广西第一产业经济值的插值多项式 $L_2(x)$ 为：

$$L_2(x) = \sum_{k=0}^2 x_k l_k(x) \quad (3)$$

$$l_k(x) = \prod_{\substack{i=0 \\ i \neq k}}^2 \frac{x - x_i}{x_k - x_i} \quad (4)$$

将式(3)按照变量 x 进行降幂排列，可以得到以下二次多项式：

$$L_2(x) = a_0 + a_1 x + a_2 x^2 \quad (5)$$

根据二次插值基函数构造法可得：

$$L_2(x) = l_0(x) y_0 + l_1(x) y_1 + l_2(x) y_2$$

此时基函数为

$$\begin{cases} l_0(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \\ l_1(x) = \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} \\ l_2(x) = \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \end{cases}$$

数据来源于广西壮族自治区统计局网站,以 2022 年广西第一产业前三季度经济值数据为训练集构建模型,以第四季度数据为测试集。具体数据如表 1 所示:

Table 1. The total value of Guangxi primary industry economy in the first three quarters of 2022
表 1. 2022 年前三季度广西第一产业经济总值

季度	经济总值
2022 年第一季度	477.7
2022 年第二季度	637.7
2022 年第三季度	1207.63
2022 年第四季度	1946.78

则 (x_0, y_0) , (x_1, y_1) , (x_2, y_2) 分别为 $(1, 477.70)$, $(2, 637.70)$, $(3, 1207.63)$, 解得 2022 年广西第一产业经济值的插值多项式 $L_2(x)$ 为:

$$L_2(x) = 204.965e^2 - 454.895e + 727.63$$

可知

$$\begin{cases} a_0 = 204.97 \\ a_1 = -454.90 \\ a_2 = 727.63 \end{cases}$$

将第一季度的第一产业经济值除以 4, 得到 5 个插值点, 再除去前后端点, 将剩余的 3 个插值节点由小到大代入多项式 $L_2(x)$ 中, 得到对应函数值, 即为 2022 年广西第一产业 1 月、2 月和 3 月的经济值。第二季度、第三季度以此类推。

利用二次插值法得到的具体数据如表 2 所示:

Table 2. The total value of Guangxi primary industry economy from January to September 2022
表 2. 2022 年 1 月至 9 月广西第一产业经济总值

月份	经济总值
1 月	166.76
2 月	154.17
3 月	156.77
4 月	174.55
5 月	207.51
6 月	255.65

Continued

7月	318.97
8月	397.48
9月	491.17

3. 建立模型

3.1. 改进灰色预测模型

灰色系统即包含已知信息又包含未知信息的系统，它在一定范围内变化，并与与时间有关，系统内各因素间的关系具有不确定性，灰色预测就是对该系统进行预测。例如，在我国整个宏观经济就是一个灰色系统，并且经济变量的增长隐含一定的指数变化趋势，它即受到确定性因素的影响，如国家宏观经济政策，又受到不确定因素的影响，如市场经济政策，因此符合灰色预测的建模条件。灰色预测通过将原始序列转换为可以发现系统变化规律的生成序列，产生具有较强规律性的数据序列，建立微分方程模型，以预测未来某一时刻的特征量，以及达到某一特征量的时间，最后进行关联分析，以识别系统各因素发展趋势之间的相异程度。根据灰色理论建立的模型为GM(1,1)模型。因原始时间序列具有随机性，为了弱化这一性质，可以使用累加的方式对原始序列进行处理，从而得到一组增长趋势明显的时间序列，然后基于此时间序列建立灰色预测模型，生成累加预测值，最后利用累减的方法进行反向计算，恢复原始时间序列，获得最终的原始预测值。GM(1,1)建模过程如下。

为了弱化原始时间序列的不稳定性，需要向灰色模型提供信息，因此在建立预测模型之前首先对原始时间序列进行预处理，经过预处理后的时间序列即被称为生成列。本文采用累积的方法进行数据处理。累加的原则是：将原始序列的第一个数据作为生成列的第一个数据，将原始序列的第二个数据与前一个数据相加，作为生成列的第二个数据，同样原始序列第三个数据与前一个数据相加，作为生成列的第三个数据，依次类推，将原始序列的最后一个数据与之前的所有数据相加，最终得到生成列，也称为累加列。

记原始时间序列为

$$p^{(0)} = \{p^{(0)}(1), p^{(0)}(2), p^{(0)}(3), \dots, p^{(0)}(n)\}$$

记累加列为

$$p^{(1)} = \{p^{(1)}(1), p^{(1)}(2), p^{(1)}(3), \dots, p^{(1)}(n)\}$$

式中的

$$p^{(1)}(k) = \sum_{i=1}^k p^{(0)}(i) = p^{(1)}(k-1) + p^{(0)}(k)$$

其中 $k = 1, 2, 3, \dots, n$ 。

根据GM(1,1)的理论定义，GM(1,1)的一阶线性微分方程如下：

$$\frac{dp^{(1)}}{dt} + \alpha p^{(1)} = \mu$$

式中， α 称为发展灰度， μ 称为内生控制灰数。

对 GM(1,1)线性微分方程求解

$$p^{(1)} = \frac{\mu}{\alpha} + Ce^{-\alpha t}$$

将上式离散化得

$$\hat{p}^{(1)}(k+1) = \frac{\mu}{\alpha} + Ce^{-\alpha k} \quad \#k = 0, 1, 2, \dots, n$$

其中 k 为时间。为求解未知量 α 和 μ ，记累加列 $p^{(1)}$ 的紧邻生成列为

$$q(k) = \frac{1}{2} [p^{(1)}(k) + p^{(1)}(k+1)] \quad \#k = 0, 1, 2, \dots, n$$

设 $\hat{\phi}$ 为代估向量， $\hat{\phi} = \begin{pmatrix} \alpha \\ \mu \end{pmatrix}$ ，利用最小二乘法求解，可得：

$$\hat{\phi} = (B^T B)^{-1} B^T Y_n$$

其中：

$$B = \begin{bmatrix} -q(2) & 1 \\ -q(3) & 1 \\ \vdots & \vdots \\ -q(n) & 1 \end{bmatrix}$$

$$Y_n = (p^{(0)}(2), p^{(0)}(3), \dots, p^{(0)}(n))^T$$

利用最小二乘法求解方程，解得未知数 α 与 μ 。为解出常数 C ，需给定条件，令

$$\hat{p}^{(1)}(1) = p^{(1)}(1) = p^{(0)}(1), \quad \text{当 } k=0 \text{ 时, } C = p^{(0)}(1) - \frac{\mu}{\alpha}.$$

则

$$\hat{p}^{(1)}(k+1) = \frac{\mu}{\alpha} + \left(p^{(0)}(1) - \frac{\mu}{\alpha} \right) e^{-\alpha k} \quad \#k = 0, 1, 2, \dots, n$$

这是我们熟悉的 GM(1,1)的预测值 $\hat{p}^{(1)}(k+1)$ 的表达式。得到这样的预测曲线 $\hat{p}^{(1)}(k+1)$ 在坐标平面上必然经过原始序列第一点，然而从最小二乘法原理来看，预测曲线并不一定需要通过第一个数据点。所以将 $\hat{p}^{(1)}(1) = p^{(1)}(1) = p^{(0)}(1)$ 作为已知条件的理论依据并不可靠。另外应考虑到 $p^{(1)}(1)$ 是一个最旧的数据，蕴含的信息量较小，并且不是通过累加生成得到的，规律性不强。因此，为了改进这个问题，我们有必要更新传统的以 $\hat{p}^{(1)}(1) = p^{(1)}(1) = p^{(0)}(1)$ 为已知条件的建模方法，考虑选用其他数据。以第 m 个数据点作为已知条件，然后再进行预测，这种方法有可能提供更准确的预测结果。例如以 $\hat{p}^{(1)}(1) = p^{(1)}(m) (m=2, 3, 4, \dots, n)$ 作为已知条件，从而解得新的预测公式。

所以，更新的公式为

$$\hat{p}^{(1)}(k+1) = \frac{\mu}{\alpha} + \left(p^{(1)}(m) - \frac{\mu}{\alpha} \right) e^{\alpha(m-k-1)}$$

其中 $k=0, 1, 2, \dots, n$ ， $m=2, 3, 4, \dots, n$ ，这里的 m 可以根据实际情况从 $2, 3, 4, \dots, n$ 中选择，形成新公式，作为原预测公式的优化，如果取 $m=1$ 时，与原预测公式相同。

3.2. 模型检验

本文中, 检验灰色预测模型的方法有残差检验、关联度检验和后验差检验。

3.2.1. 残差检验

首先按预测模型计算 $\hat{p}^{(1)}(k)$, 并将 $\hat{p}^{(1)}(k)$ 累减生成 $\hat{p}^{(0)}(k)$, 然后计算原始序列 $p^{(0)}(k)$ 与 $\hat{p}^{(0)}(k)$ 的绝对误差序列 $\Delta^{(0)}(i)$ 及相对误差序列 $\Phi(i)$ 。若相对误差小于 0.5%, 代表模型精确度较高。

$$\Delta^{(0)}(i) = |p^{(0)}(k) - \hat{p}^{(0)}(k)| \# k = 1, 2, 3, \dots, n$$

$$\Phi(i) = \frac{\Delta^{(0)}(k)}{p^{(0)}(k)} \times 100\% \# k = 1, 2, 3, \dots, n$$

3.2.2. 关联度检验

在客观世界中, 因素之间存在许多联系, 有关系密切的因素, 关系不密切的因素, 还有许多因素之间的关系时灰色的。要找到主要矛盾, 发现主要特征和主要关系, 就利用到关联度分析方法。关联度分析是计算系统中各因素关系密切度的方法, 在得出计算关联度之前, 还需要获知关联系数。本文中关联系数 $\eta(k)$ 是反映原始预测值 $\hat{p}^{(0)}(k)$ 与原始值 $p^{(0)}(k)$ 在第 k 点的关联程度, 计算公式如下。

设:

$$\hat{p}^{(0)}(k) = \{\hat{p}^{(0)}(1), \hat{p}^{(0)}(2), \dots, \hat{p}^{(0)}(n)\}$$

$$p^{(0)}(k) = \{p^{(0)}(1), p^{(0)}(2), \dots, p^{(0)}(n)\}$$

关联系数定义为:

$$\eta(k) = \frac{\min \min |\hat{p}^{(0)}(k) - p^{(0)}(k)| + \rho \max \max |\hat{p}^{(0)}(k) - p^{(0)}(k)|}{|\hat{p}^{(0)}(k) - p^{(0)}(k)| + \rho \max \max |\hat{p}^{(0)}(k) - p^{(0)}(k)|}$$

其中

- 1) $|\hat{p}^{(0)}(k) - p^{(0)}(k)|$ 是第 k 点预测值与原始值的绝对误差;
- 2) $\min \min |\hat{p}^{(0)}(k) - p^{(0)}(k)|$ 是两级最小差;
- 3) $\max \max |\hat{p}^{(0)}(k) - p^{(0)}(k)|$ 是两级最大差;
- 4) ρ 称为分辨率, 一般取 $\rho = 0.5$;

关联度是参考原始预测值与各原始值在各个时刻的关联系数的平均值, 计算公式如下:

$$r = \frac{1}{n} \sum_{k=1}^n \eta(k)$$

当 $\rho = 0.5$ 时, 关联度 $r > 0.6$ 便满意了。关联度越接近 1, 关联系数越强。

3.2.3. 后验差检验

- 1) 计算原始序列的标准差:

$$S_1 = \sqrt{\frac{\sum [p^{(0)}(k) - \bar{p}^{(0)}]^2}{n-1}}$$

2) 计算绝对误差序列的标准差:

$$S_2 = \sqrt{\frac{\sum [\Delta^{(0)}(k) - \bar{\Delta}^{(0)}]^2}{n-1}}$$

3) 计算方差比:

$$D = \frac{S_2}{S_1}$$

4) 计算小误差概率:

$$G = g \left\{ \left| \Delta^{(0)}(k) - \bar{\Delta}^{(0)} \right| < 0.6745S_1 \right\}$$

若 $S_0 = 0.6745S_1$, 则 $G = g \left\{ \left| \Delta^{(0)}(k) - \bar{\Delta}^{(0)} \right| < S_0 \right\}$ 。

后验差精度检验等级参照表如表 3:

Table 3. Precision inspection level reference table

表 3. 精度检验等级参照表

D 值	G 值	等级
$D \leq 0.35$	$0.95 \leq p$	好
$0.35 < D \leq 0.5$	$0.8 \leq p < 0.95$	合格
$0.5 < D \leq 0.65$	$0.7 \leq p < 0.8$	勉强合格
$0.65 < D$	$p < 0.7$	不合格

文中, 以上检验的 n 为序列长度, 即数据个数。若残差检验、关联度检验和后验差检验都能通过, 则说明所建模型精度较好, 可以用其进行预测。

4. 广西第一产业经济发展预测和分析

4.1. 模型预测

首先, 用二次插值法求出的 2022 年 1 月至 9 月广西第一产业经济总值数据, 利用 R 软件编程对数据进行计算, 求得模型的系统发展系数 $\alpha = -0.1934$, 内生控制灰数 $\mu = 77.67854$ 。经过计算发现, 在预测公式中, 当 $m = 7$ 、 $p^{(1)}(7) = 1434.38$ 时模型精度最高, 此时预测模型为:

$$\hat{p}^{(1)}(k+1) = \frac{77.6785}{-0.1934} + \left[1434.38 - \frac{77.6785}{-0.1934} \right] e^{-0.1934(m-k-1)}$$

再递减排还原得 2022 年广西第一产业经济预测值:

$$\hat{p}^{(0)}(k) = \hat{p}^{(1)}(k) - \hat{p}^{(1)}(k-1)$$

利用上述表达式预测广西第一产业经济, 得预测收入及相对误差, 如表 4 所示。在模型所得指标中, 数值的相对误差最为关键, 具有较强的参考意义。就第一产业经济自身特性而言, 经济预测值相对误差小于 5% 时, 可视为高精度预测, 相对误差为 5% 至 10% 时, 可视为较好预测, 由表 4 可以看出相对误差较小的是 7 月、8 月和 9 月, 分别为 1.21%、1.45% 和 3.24, 因为这三个月在第三季度, 数据

较新所以误差较小；较相对误差较大的为 2 月、5 月和 3 月，分别为 20.37%、5.68% 和 4.99%，其中 2 月数据为异常值。

Table 4. Model fitting results table
表 4. 模型拟合结果表

月份	第一产业经济总值	预测值	相对误差
2022 年 1 月	166.76	173.75	4.19%
2022 年 2 月	154.17	122.76	20.37%
2022 年 3 月	156.77	148.95	4.99%
2022 年 4 月	174.55	180.73	3.54%
2022 年 5 月	207.51	219.29	5.68%
2022 年 6 月	255.65	266.07	4.08%
2022 年 7 月	318.97	322.83	1.21%
2022 年 8 月	397.48	391.70	1.45%
2022 年 9 月	491.17	475.27	3.24%

Table 5. The results of correlation degree test and posterior difference test
表 5. 关联度检验、后验差检验结果

相对精度	关联度	<i>D</i> 值	<i>G</i> 值
94.58%	0.68	0.0049	1

由拟合结果表 5 可知，模型精度为 94.58%，模型关联度为 $0.68 > 0.6$ ，模型的精确度较高。模型的精度检验等级参照表如表 3 所示，后验差检验中，*D* 值为 $0.0049 < 0.35$ 且小误差概率 *G* 值为 $1 > 0.95$ ，可知该模型为优模型，可用于预测未来广西第一产业经济总值数据。

4.2. 模型分析

本文对 2022 年 10 月至 12 月广西第一产业经济总值进行预测，结果如表 6 所示；利用改进后的 GM(1,1) 预测模型对实际值和预测值进行拟合，结果如图 1。预测结果显示，2022 年 10 月广西第一产业经济预测值为 576.66 亿元，11 月预测值为 699.68 亿元，12 月预测值为 838.641 亿元。模型测算的广西第一产业第四季度经济总值为 2125.29 亿元，测试集中的真实值为 1946.78 亿元，整体走势符合实际。

Table 6. Forecast of Guangxi primary industry economic value from October to December 2022
表 6. 2022 年 10 月至 12 月广西第一产业经济总值预测

月份	预测值(单位: 亿元)
2022 年 10 月	576.66
2022 年 11 月	699.68
2022 年 12 月	848.95

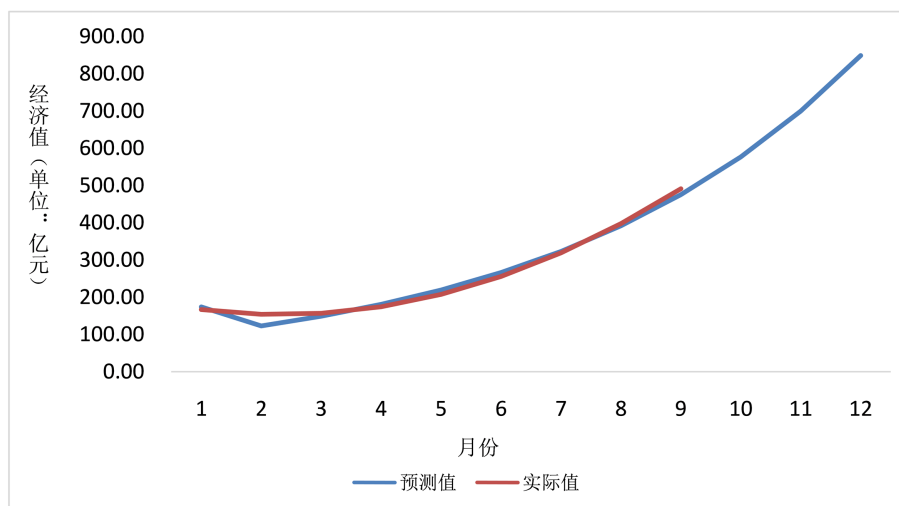


Figure 1. Improved GM(1,1) prediction model fitting diagram
图 1. 改进的 GM(1,1) 预测模型拟合图

从模型拟合图可以看出, 预测数据呈指数增长, 在 2 月份有所回落, 广西第一产业经济总体增长趋势显著, 这可为地方政府制定发展农林牧渔业相关政策提供科学有效的模型支撑。新冠肺炎疫情发生以来, 对于农业经济也产生较大的冲击, 对广西第一产业经济的影响也不可避免, 导致预测的数据会有一些的偏差, 如 2 月份的数据误差较大。

4.3. 模型适用性

改进的灰色 GM(1,1) 模型具有广泛的适用性, 能够应对各种预测和分析场景, 特别是在数据不足、信息不完全、变化规律不明显的情况下, 这种模型显得尤为重要。在经济预测方面, 该模型可以帮助预测和分析多种重要经济指标, 如 GDP、就业率、通货膨胀率和财政收入等, 为经济决策提供可靠的参考依据。在农业预测方面, 无论是预测农产品的产量、种植面积还是农业机械化水平, 该模型都能提供准确的趋势预测, 对于指导农业生产和政策制定具有重要作用。此外, 该模型在环保、交通、能源、公共卫生和教育等领域的预测和分析任务中也能发挥出色的效果。无论是预测空气质量、交通流量、能源需求, 还是疾病的发展趋势和学生入学率, 改进的灰色 GM(1,1) 模型都能提供相对准确的预测结果。此模型的主要优势在于其能够揭示系统的发展趋势, 从而为决策提供科学的参考。这使得它在数据量较少、不容易收集到完整数据或者数据变化趋势不明显的情况下, 都能发挥其独特的价值。总的来说, 改进的灰色 GM(1,1) 模型的适用范围广泛, 为不同领域的预测分析提供了一种简单且高效的工具, 其在未来的各种预测任务中的应用前景十分广阔。

4.4. 模型对比及优缺点

未改进的 GM(1,1) 预测模型对实际值和预测值进行拟合, 结果如图 2。虽然前期拟合效果较好, 但后期逐渐产生误差, 不能较为精确的把握数据未来的发展趋势。

改进的灰色 GM(1,1) 模型主要优势有适用范围非常广泛, 适用于预测和分析数据不足、信息不完全、变化规律不明显的系统, 并且允许模型有数据量较少、不容易收集到完整数据或者数据变化趋势不明显的情况; 其次是趋势预测力强, 这种模型注重的是发展趋势, 而非精确数值, 使其在长期趋势预测中展现出强大的实力; 还有改进后的模型提高了预测精度, 通过使用其他数据点作为已知条件, 改进的 GM(1,1) 模型能够提供更精确的预测, 显著提升了预测的精度和效果。

虽然经过检验表明改进后的模型精度也有所提高,但仍存在局限性。如第一是数据波动较大或变化趋势不明显的情况下,预测精度可能仍然受限;第二是此类模型主要适用于处理线性或近线性问题,对于非线性问题,其预测效果可能不理想;第三是模型需要对数据进行适当的预处理,如数据频率的转换,这可能增加了数据处理的复杂性。

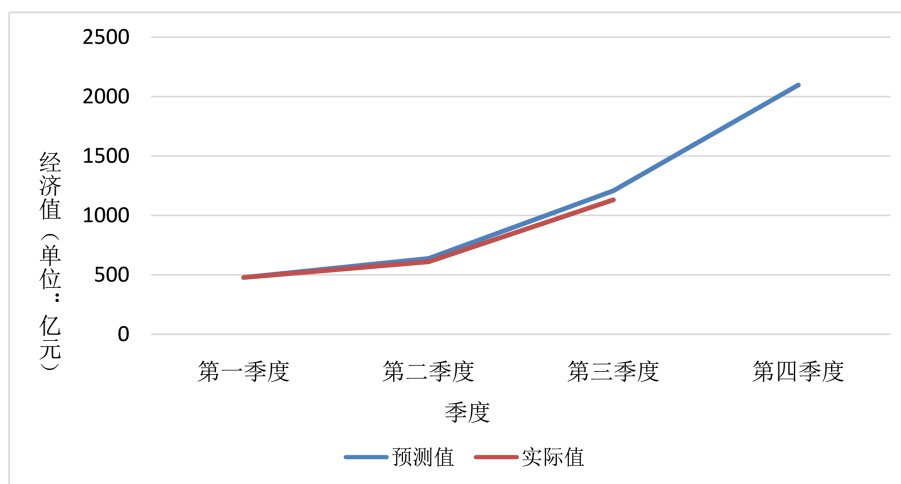


Figure 2. Unimproved GM(1,1) prediction model fitting diagram

图 2. 未改进的 GM(1,1) 预测模型拟合图

5. 结论

广西是传统农业大区,第一产业经济总值则是经济基本盘的具体表现。党中央对三农工作的高度重视,表现在举全党之力全面推进乡村振兴,加快农业农村现代化,加快建设农业强国,建设宜居宜业的美丽乡村。第一产业经济作为和农业息息相关的数值,是推进农业现代化和增加农民收入的聚焦关键点。在深刻把握习总书记“稳字当先”工作要求的基础上,通过计算可知广西第一产业经济总值整体呈上升态势。因此,本文深入研究广西第一产业经济总值变化态势,利用 2022 年前三季度已知数据,建立广西第一产业经济总值的灰色预测 GM(1,1)模型,该模型通过了残差检验、关联度检验与后验差检验,模型精度较高,对广西农村人均可支配收入预测研究中有实际意义。综上所述,虽然新冠肺炎疫情发生以来,广西第一产业经济发展受到了某些方面的影响,导致预测的数据存在不稳定性;但通过把农业大数据与预测结果结合,可为相关部门提供农业方面的信息和资源,为广大农户进行农业生产提供重要参考和借鉴。广西政府应以实现建设农业强省为目标,采取有利于农业增收的有效措施,再要结合广西农业特色,立足当地农耕情况,在提高一产经济值的同时,持续优化升级农业发展品质,为推动农业经济可持续发展奠定坚实的基础。

参考文献

- [1] 谭学瑞, 邓聚龙. 灰色关联分析: 多因素统计分析新方法[J]. 统计研究, 1995(3): 46-48. <https://doi.org/10.19343/j.cnki.11-1302/c.1995.03.011>
- [2] 李云贵, 李清富, 赵国藩. 灰色 GM(1,1) 预测模型的改进[J]. 系统工程, 1992(6): 27-31.
- [3] 吴晓明, 关蓬莱. 多维自回归滑动平均模型研究与应用[J]. 辽宁大学学报(自然科学版), 2002(1): 22-27.
- [4] 王妍, 高强, 金炜博. 基于灰色系统理论的山东省农业产业结构优化研究[J]. 山东农业科学, 2015, 47(3): 144-147+153. <https://doi.org/10.14083/j.issn.1001-4942.2015.03.037>
- [5] 李勇, 王毅彰. 河南省第一产业增加值的关联分析与灰色预测[J]. 安徽农业科学, 2009, 37(32):

- 16004-16005+16054. <https://doi.org/10.13989/j.cnki.0517-6611.2009.32.185>
- [6] 凌立文, 陈善盈, 张大斌. 考虑层级特性的第一产业总产值预测及修正[J]. 系统工程学报, 2019, 34(3): 289-300+421. <https://doi.org/10.13383/j.cnki.jse.2019.03.001>
- [7] 张仕清. 第一产业产值预测模型及其结构优化分析[J]. 数学的实践与认识, 2012, 42(1): 68-73.
- [8] 杨国华, 颜艳, 杨慧中. GM(1,1)灰色预测模型的改进与应用[J]. 南京理工大学学报, 2020, 44(5): 575-582. <https://doi.org/10.14177/j.cnki.32-1397n.2020.44.05.010>
- [9] 田梓辰, 刘淼. 基于改进灰色 GM(1,1)模型的 GDP 预测实证[J]. 统计与决策, 2018, 34(11): 83-85. <https://doi.org/10.13546/j.cnki.tjyjc.2018.11.019>
- [10] 张大海, 江世芳, 史开泉. 灰色预测公式的理论缺陷及改进[J]. 系统工程理论与实践, 2002(8): 140-142.